Computational Methods of Optimization
First Midterm(20th Sep, 2016)

Answer all questions Time: 90 minutes

September 19, 2016

1. Consider minimization of the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as follows

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top A\mathbf{x} - b^\top \mathbf{x} + c$$

$$A = \begin{bmatrix} 1 & a \\ -a & 2 \end{bmatrix}, b = [1, 2], c = 1$$

The value of a is not known.

(a) Answer true or false. All questions carry 1 mark

i. f is not in C^2.

ii. f does not have a global minimum.

iii. f is convex

iv. It has a unique local minimum

v. The set $S = \{\mathbf{x} = [x_1, x_2]^\top | f(\mathbf{x}) \leq 1\}$ satisfies

$$S \subset \left\{ \mathbf{x} \,\middle|\, \frac{1}{2}(x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \right\}$$

(b) The following questions carry 2 marks

i. Compute the critical points of f?

ii. Compute the eigenvalues of the Hessian of f at an arbitrary point $\mathbf{x}$

iii. Compute the Lipschitz constant of f? Recall that Lipschitz constant is defined as a constant, L, such that

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|$$

holds for all points $\mathbf{x}, \mathbf{y}$ in the domain of f.

iv. Consider the following update $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha\nabla f(\mathbf{x}^{(k)})$. The stepsize α is constant across iterations. Find an interval (α_1, α_2) such that for any α in that interval $f(\mathbf{x}^{(k+1)}) \leq f(\mathbf{x}^{(k)}), k \geq 0$ is guaranteed?

v. If $\alpha = \frac{1}{L}$ determine N such that

$$f(\mathbf{x}^{(k)}) \leq -0.49, \qquad \text{for all } k \geq N$$

Assume that $\mathbf{x}^{(0)} = 0$.

2. Consider f given in Question 1. Suppose instead of constant step-size we do exact line search. For such a choice show that

$$9f(\mathbf{x}^{(k+1)}) \leq f(\mathbf{x}^{(k)}) - 4$$

Determine N, as defined in Question 1(b)v 5 marks

3. Let $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^n$ be two fixed vectors. Let α^* and f^* be defined as follows

$$f^* = min_{\alpha \in \mathbb{R}} f(\alpha) = f(\alpha^*) \qquad f(\alpha) = \frac{1}{2}\|\mathbf{u} - \alpha\mathbf{v}\|^2$$

(a) Find f^*, α^*. Justify your answer. 3 marks

(b) State and prove Cauchy Schwartz Inequality. 2 marks

(c) Using Cauchy Schwartz inequality solve the following problem

$$max_{\mathbf{u}} \mathbf{u}^\top A\mathbf{u} \qquad \|\mathbf{u}\| \leq 1$$

Assume A is a $d \times d$ square symmetric matrix and $\mathbf{u} \in \mathbb{R}^d$. Report the optimal value and the optimal $\mathbf{u}$. 5 marks

(d) Using the previous question solve the following problem

$$f^* = max_{\mathbf{x} \neq 0} f(\mathbf{x}) \qquad f(\mathbf{x}) = \frac{\mathbf{x}^\top A\mathbf{x}}{\|\mathbf{x}\|^2}$$

and report f^*. Justify your answer. 5 marks

4. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, not in C^1. Let $\mathbf{x}_1, \mathbf{x}_2$ be two local minima of f

(a) Define local minimum and global minimum of f? 2 marks

(b) Using the definitions prove or disprove the following statement. Global minimum of f is attained at $\mathbf{x}_3 = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$. 8 marks

5. Let $f(x) = \frac{1}{2}x^\top Qx - b^\top x + c$ be defined for all $x \in \mathbb{R}^d$, where $Q \in \mathbb{R}^{d\times d}, b \in \mathbb{R}^d, c \in \mathbb{R}$. It is given that Q is symmetric and positive definite. Let u_k be the steepest descent direction.

(a) Show that at a point x_k the function $h(\alpha) = f(x_k + \alpha u_k)$ can be written as $h(\alpha) = a_0(\alpha - \bar{\alpha}_k)^2 + a_1$. Explicitly state a_0, a_1 and $\bar{\alpha}$ in terms of Q, b, x_k and c. 2 marks

(b) Show that the set of acceptable stepsizes, which satisfies Goldstein conditions, can be described by the set $\left\{\alpha \middle| \left|\frac{\alpha}{\bar{\alpha}_k} - 1\right| \leq s_k\right\}$. What is the relationship between s_k and the parameter used in Goldstein condition. 2 marks

(c) Is there any advantage in using Wolfe conditons for this problem? Give reasons? 1 marks

Computational Methods of Optimization

~~First~~ Second Midterm(26th Oct, 2016)

Answer all questions Time: 90 minutes

October 26, 2016

1. (a) State and derive the rank-2 Quasi-Newton update? 5 marks

   (b) Prove that the update for a Quadratic convex function with full rank Hessian always yields a positive semidefinite matrix whenever exact line search is used. 10 marks

2. Bunty and Babli were arguing over the following problem

$$min_{x_1 \in \mathbb{R}, x_2 \in \mathbb{R}} x_1^2 + x_2^2 \text{ subject to } x_1^2 = (x_2 - 3)^5 \quad (\mathcal{P})$$

   (a) Bunty substitutes x_1^2 in the objective by $(x_2 - 3)^5$ and transforms $(\mathcal{P})$ into the following unconstrained problem

$$min_{x_2} (x_2 - 3)^5 + x_2^2 \quad (\mathcal{Q})$$

   ' The objective function is not bounded from below and global minimum does not exist. Hence Bunty concludes that global minimum of $(\mathcal{P})$ does not exist. Do you agree with Bunty's views? Give reasons? 5 marks

   (b) Babli disagrees with Bunty and tells him that global minimum of $(\mathcal{P})$ exists. She thinks that Bunty has made a mistake but to correct it he needs to add a constraint to $(\mathcal{Q})$ and then he will be able to find the global minimum of $(\mathcal{P})$. If you disagree with Babli give reasons? If you agree with Babli find the constraint that needs to be added to $(\mathcal{Q})$ and then find the global optimal value of $(\mathcal{P})$ as suggested by Babli? 5 marks

3. Let x^* be a point with primal objective value p^* and λ^* be dual optimal with dual objectve value d^*. If $p^* = d^*$ show that x^* must be primal optimal and d^* must be the dual optimal. value 5 marks

4. Consider the following problem

$$min_{\mathbf{x}\in\mathbb{R}^d} \mathbf{x}^\top g + \frac{1}{2}\mathbf{x}^\top B\mathbf{x}$$

$$\mathbf{x}^\top \mathbf{x} \leq 1$$

where B is a symmetric $d \times d$ matrix while g is a $d \times 1$ vector. B has both positive and negative eigenvalues.

(a) State the KKT conditions of the problem. Compute the dual function, $h(\lambda)$, and state the dual optimization problem. 5 marks

(b) For the case $g = 0$, solve the dual optimization problem. Show that strong duality holds and then solve the primal problem 5 marks

(c) Solve the dual problem for the following setting

$B \in \mathbb{R}^{2\times 2}$ and is a diagonal matrix with $B_{11} = -B_{22} = 1.$

and $g = [1, 1]^\top$. Does strong duality hold for this setting. Give reasons. 5 marks

5. Solve the following problem

$$max_{\mathbf{x}\in\mathbb{R}^d} - \sum_{i=1}^{d} x_i \log x_i \;\; \sum_{i=1}^{d} x_i = 1, x_i \geq 0$$

5 marks

Computational Methods of Optimization

~~First Midterm~~(14th Dec, 2016)

Answer all questions Time: 180 minutes

December 13, 2016

1. A function $f : \mathbb{R}^d \to \mathbb{R}$ satisfies the following property

$$f(\mathbf{y}) \leq f(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^\top \nabla f(\mathbf{x}) + \frac{\mu}{2}\|\mathbf{x} - \mathbf{y}\|^2$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. Consider an iterative algorithm which at iteration k takes input $\mathbf{x}^{(k-1)}$ and outputs $\mathbf{x}^{(k)}$ such that

$$f(\mathbf{x}^{(k)}) \leq f\left(\mathbf{x}^{(k-1)} - \frac{1}{\mu}\nabla f(\mathbf{x}^{(k-1)})\right)$$

(a) Assuming that the algorithm starts at any arbitrary $\mathbf{x}^{(0)}$ show that there exists a constant A such that

$$min_{1 \leq k \leq T}\|\nabla f(\mathbf{x}^{(k-1)})\| \leq \frac{A}{\sqrt{T}}$$

5 marks

(b) Suppose we do steepest descent with constant stepsize,

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha \nabla f(\mathbf{x}^{(k)})$$

For what values of α could we derive a constant A as in the previous question. 5 marks

2. Consider the following statement.

The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^d$ are linearly dependent and also $Q-$conjugate, where Q is a real symmetric $d \times d$ matrix. Assume $n < d$.

Under what conditions on Q is the above statement true? 5 marks

3. State the conjugate gradient(CG) algorithm? 5 marks

4. Let $Q = \mathbf{u}\mathbf{u}^\top + 2\mathbf{v}\mathbf{v}^\top$ where $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$ are vectors with $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ and $\mathbf{u}^\top\mathbf{v} = 0$. Consider the unconstrained minimization of the function $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} + \mathbf{w}^\top\mathbf{x} + c$ where $\mathbf{w}^\top\mathbf{u} = \mathbf{w}^\top\mathbf{v} = 0$

   (a) Consider applying the CG algorithm to the problem. Compute the number of iterations the CG algorithm would take to converge starting from $\mathbf{x}^{(0)} = 0$. 5 marks

   (b) How does your answer change if $\mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v}$ where α, β are scalars such that $\alpha^2 + \beta^2 = 1$ 5 marks

5. Consider minimizing $f : \mathbb{R}^d \to \mathbb{R}$ over a set $C \subset \mathbb{R}^d$. Suppose you were able to find $\mathbf{x} \in C$ such that $f(\mathbf{x}) = 1$ and a dual feasible point λ such that the dual objective function value is 0.99. Can you estimate how close is $f(\mathbf{x})$ to the optimal objective function value. Suppose now that the function f is μ-strongly convex and C is also a convex set. Can you estimate how close is $\mathbf{x}$ to the optimal point. A function f is said to be $\mu-$strongly convex on C if there exists $\mu > 0$ such that

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top(\mathbf{y} - \mathbf{x}) + \mu\|\mathbf{y} - \mathbf{x}\|^2$$

holds for all $\mathbf{x}, \mathbf{y} \in C$. 5 marks

6. Consider the problem of minimizing the following function

$$min_{\mathbf{x}\in\mathbb{R}^2} e^{x_2} + \frac{1}{2}x_1^2 - x_2$$

$$\text{subject to } x_1^2 + x_2^2 \le 3, \quad a^\top x \ge 2$$

Consider the point $\mathbf{x}(0) = [1, 1]^\top$. For what values of a is $\mathbf{x}^{(0)}$ a local minimum. 5 marks

7. Consider the following problem. Let

$$\begin{aligned} min_{\mathbf{x}\in\mathbb{R}^2} \quad & 4x_1 - 2x_2 \\ s.t \quad & 2x_1 + 3x_2 \ge 6 \\ & x_1 + x_2 \ge 1 \\ & \mathbf{x} \ge 0 \end{aligned}$$

   (a) Derive the Dual optimization problem 5 marks
   (b) Compute a Dual feasible point 2 marks
   (c) Repeat the above steps by replacing the constraint $2x_1 + 3x_2 \ge 6$ with $2x_1 + 3x_2 \le 6$. Compute the optimal dual and primal objective function value in both cases. 8 marks

8. Use the penalty function method to minimize

$$min_{\mathbf{x}\in\mathbb{R}^2} x_1^2 + x_2^2 \quad x_1 + x_2 \ge 1$$

10 marks

9. Consider the gradient projection method to solve the following problem

$$min_{\mathbf{x}\in\mathbb{R}^d} \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} - \mathbf{h}^\top\mathbf{x} + c$$

$$\text{subject to } a_i \le x_i \le b_i \; i \in \{1, \dots, d\}$$

Assume Q is positive definite and symmetric. Compute one iteration of the algorithm starting from $\mathbf{x}^{(0)} = 0$. Justify your answer. 10 marks